

Increased Frequency of N-Region Insertion in a Murine Pre-B-Cell Line Infected with a Terminal Deoxynucleotidyl Transferase Retroviral Expression Vector

NATHANIEL R. LANDAU,[†] DAVID G. SCHATZ, MARGARET ROSA, AND DAVID BALTIMORE*

Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, and Biology Department, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received 4 May 1987/Accepted 5 June 1987

The role of terminal deoxynucleotidyl transferase (TdT) in the insertion of N regions into the junctional sites of immunoglobulin genes was investigated. Pre-B-cell lines capable of continuous rearrangement of immunoglobulin light-chain genes and differing only in the presence or apparent absence of TdT were derived by infecting cells with a TdT retroviral expression vector or a control vector. The cell lines were then superinfected with a retrovirus-based artificial immunoglobulin gene rearrangement substrate. The substrate was allowed to rearrange in the cell lines and the rearranged proviruses were rescued from the cell lines. Nucleotide sequence analysis of the V-J junctions of the proviral rearranged genes showed a fivefold greater frequency of N-region insertion in proviruses rescued from the TdT⁺ cell lines than in those rescued from the TdT⁻ cell lines, so that at least 50% of the rearrangements that occurred in the presence of TdT had N regions. It is thus evident that TdT can stimulate N-region insertion, and the enzyme is presumably directly responsible for adding nucleotides at V-J and other immunoglobulin and T-cell receptor gene junctions.

Terminal deoxynucleotidyl transferase (TdT) is a unique DNA polymerase that without template direction catalyzes the addition of deoxyribonucleotides onto the 3'-hydroxyl ends of DNA primers (5, 15). It is present in the immature fraction of thymocytes, in a small fraction of bone marrow cells, in transformed pre-B- and pre-T-cell lines, and in leukemic cells (8, 10, 18, 27, 35). The enzyme is composed of a single polypeptide chain of molecular weight 58,000 (6, 29, 36). cDNA clones have been isolated representing mouse (20), human (31), and bovine (17) TdT.

Initially, it was thought that TdT might be responsible for the generalized somatic mutation seen throughout the variable regions of immunoglobulin genes (2). Later, it became evident that this somatic mutation process takes place in the later stages of B-lymphoid development (13), when TdT is no longer present (18), and somatic mutation probably does not occur at all in T-lymphoid cells (9). Another suggestion for the function of TdT is that it might be responsible for insertion of small numbers of nucleotides of largely random sequence at the DNA joints formed during the construction of immunoglobulin genes from germ line gene fragments. The sequences of an odd rearrangement in a heavy-chain gene first suggested this notion (1), and further studies of joints formed in cells containing or lacking TdT have substantiated it (4, 11, 39). Sequences not evident in surrounding germ line DNA have been observed at V-D and D-J junctions of both immunoglobulin heavy chains (19) and T-cell receptor genes (3) but not at the V-J junctions of immunoglobulin light-chain genes (34). An explicit proposal is that these nucleotides or nucleotide regions (N regions) result from the polymerization of random nucleotides by TdT onto free 3' ends of DNA exposed during gene rearrangement (1).

To test this idea of TdT function directly we have chosen to insert a functional TdT gene into a cell line that rearranges κ light-chain genes but ordinarily lacks TdT. Such a cell line is the Abelson murine leukemia virus-transformed line PD31 (23). We introduced an expressed, exogenous TdT gene through a retroviral vector and then superinfected the cells with a virus containing κ immunoglobulin gene segments able to rearrange (21, 22). If TdT was capable of synthesizing N-regions, we expected to find that immunoglobulin genes which had rearranged in the presence of TdT would show an increased frequency of N regions as compared with those that had rearranged in cells containing little or no TdT. This expectation was fulfilled.

MATERIALS AND METHODS

Construction of vectors. The TdT retroviral vectors were constructed by digestion of pM-TdT or pHM-TdT (see below) with *Bam*HI and *Eco*RI, ligation to *Bam*HI linkers, and ligation to *Bam*HI-cleaved pDOL⁻. To construct the DGR retroviral vector, a cloned BALB/c J κ 1 gene segment (25), was cleaved with *Hind*III and *Hinf*I and reacted with the large subunit of DNA polymerase I to blunt the ends and ligated to *Nru*I-cleaved pBR322. The resulting plasmid was cleaved with *Sal*I and ligated to the *Sal*I-linkered J κ 1 restriction fragment. This plasmid was cleaved with *Sph*I and *Bam*HI and ligated to *Bam*HI-*Sph*I-cleaved V κ 21-C DNA (isolated from pVJG [21]). The resulting plasmid was cleaved with *Sph*I and ligated to the *Bam*HI-*Bgl*II-cleaved, *Sph*I-linked restriction fragment of pVJG containing the *gpt* gene and the 3' splice site of murine leukemia virus. A restriction fragment containing V κ 21-C, two copies of J κ 1, and the *gpt* gene was excized by cleaving with *Nhe*I-*Pvu*II, ligated to *Bgl*II linkers, and ligated to *Bam*HI-cleaved pDOL.

Cell culture. To establish virus producer cell lines, ψ -2 cells (24) were transfected with 10 μ g of plasmid by the protocol of Graham and Van der Eb (14) as modified by Parker and Stark (30). Virus was harvested 18 h after the glycerol treatment, filtered through a 0.45- μ m filter, and

* Corresponding author.

[†] Present address: Department of Microbiology and Immunology, University of California, San Francisco, CA 94143.

used to infect ψ -2 cells that had been grown for 18 h in the presence of tunicamycin (Calbiochem-Behring, La Jolla, Calif.). The tunicamycin-treated ψ -2 cells were incubated with virus for 2 h in the presence of 8 μ g of polybrene per ml. The virus was removed, and the cells were grown for 2 days in nonselective medium, at which point they were placed into medium containing 1 mg of G418 (GIBCO Laboratories, Grand Island, N.Y.) per ml. Single colonies were picked 8 to 10 days later and analyzed by filter hybridization. The titer of virus produced by the ψ -2 producer lines was determined by adding supernatant from each of the ψ -2 clones to 3T3 cells, followed by selection of the 3T3 cells in 1 mg of G418 per ml and counting the number of surviving cell clones. Lymphocytes were infected by culturing 10^5 cells for 2 days on a monolayer of ψ -2 producer cells (titer, $\approx 2 \times 10^4$ CFU/ml). Lymphocytes were removed from the ψ -2 cells by gentle aspiration and cultured in G418 (2 mg/ml) or mycophenolic acid (0.5 μ g/ml).

Rescue of proviruses. High-molecular-weight DNA (40 μ g) was cleaved to completion with *Xba*I. The DNA was electrophoresed on 0.8%, agarose and the region of the gel from 7 kilobases (kb) to 8 kb was excised. The DNA was purified by the glass powder method (38) and dissolved in 20 μ l of TE (10 mM Tris [pH 8.0], 1 mM EDTA). A sample of the DNA (5 μ l) was reacted with T4 DNA ligase in a total volume of 200 μ l at 25°C for 2 h. The DNA was extracted with water-saturated phenol and precipitated by the addition of 1 μ g of yeast tRNA, 500 μ l of ethanol, and 20 μ l of 3 M sodium acetate (pH 6.2). The DNA was incubated on dry ice for 20 min and centrifuged in a microfuge for 10 min. The pellet was washed in ethanol, dried, and dissolved in 10 μ l of TE. The DNA was used to transform 100 μ l of competent *Escherichia coli* MC1061 (transformation efficiency, 5×10^8 colonies per μ g of pUC13). The transformed bacteria were plated on Luria broth agar containing 50 μ g of kanamycin per ml.

Nucleic acid hybridization. High-molecular-weight DNA (10 μ g) was cleaved with restriction enzyme, fractionated by electrophoresis through 1.0% agarose, transferred to nitrocellulose filters and hybridized as described previously (37). Probes were labeled with [α - 32 P]dCTP by the hexamer labeling procedure (12).

Enzyme assays. TdT assays were performed as previously described (28). Briefly, cells were lysed by the addition of lysis buffer containing 0.1% Triton X-100–3 M KCl. Lysates were applied to 0.5-ml phosphocellulose columns. The columns were washed with buffer containing 0.3 M KCl–10 mM Tris (pH 8.0)–1 mM EDTA and eluted with the same buffer containing 0.5 M KCl. The column fractions were assayed for TdT activity in a buffer containing 50 mM Tris (pH 8.0)–1 mM dithiothreitol–5 μ g of oligo(dA)₁₂₋₁₈ per ml–1 μ Ci of [α - 32 P]dGTP–2 nM dGTP–0.2 mM MnCl₂, and incorporation was measured by scintillation counting of trichloroacetic acid-precipitable label.

RESULTS

TdT retroviral expression vector. A TdT gene functional in bacteria was constructed by fusing the 5' end of a mouse TdT cDNA to a human cDNA clone lacking 5' sequence; this we call HM-TdT. The protein encoded by HM-TdT contained 13 amino acid differences from the authentic human enzyme (32). We had initially planned to use a full-length mouse cDNA isolated by using a partial cDNA (20). This clone, however (referred to henceforth as M-TdT), yielded only inactive protein when tested in bacteria and has been used as a negative control. The nucleotide sequence of

M-TdT was identical to that of a previously reported murine TdT cDNA with the exception of a small number of point mutations (17). The cDNAs for the inactive M-TdT and active HM-TdT were trimmed of 5' and 3' untranslated regions and inserted into the pDOL plasmid that produces the DOL retrovirus (17a), yielding pDOL-M-TdT and pDOL-HM-TdT (Fig. 1A). In addition to TdT, these retroviruses express the *neo* gene, encoding G418 resistance, driven by the simian virus 40 promoter (Fig. 1A).

Cell lines that stably produced the DOL⁺-TdT viruses were established and those producing the highest titers of G418^r virus were used in further experiments. The structure of the proviruses in the virus producer cell lines and of 3T3 cells infected by the viruses produced by the cell lines was studied by DNA blotting analysis. This analysis showed that each line contained a single provirus whose size was identical to that of the intact DOL⁺-TdT virus and that the viruses produced represented intact copies of the retroviral expression vectors (data not shown).

Expression of TdT enzymatic activity in TdT retroviral infected cells. The ability of the retroviruses to direct the synthesis of enzymatically active TdT in mammalian cells was initially tested by assaying for the presence of TdT in virus-infected 3T3 cell clones. Lysates prepared from DOL⁺-M-TdT virus-infected cell lines showed insignificant levels of TdT activity. Each of the four DOL⁺-HM-TdT virus-infected cell clones tested showed levels of TdT activity comparable to that found in the RL δ 11 cell line (Table 1). RL δ 11 has been shown previously to synthesize a level of TdT similar to that of a normal, TdT⁺ thymocyte (10). There was no morphological difference between the TdT⁺ and TdT⁺ 3T3 cells, nor did the cell lines show a clear difference in their cell division rates. TdT expression at physiological levels is therefore not toxic to 3T3 cells.

The Abelson murine leukemia virus-transformed pre-B-cell line PD31 was then infected with the DOL⁺-M-TdT and DOL⁺-HM-TdT viruses. The cells were separated into 12 groups and grown in the presence of G418 to select for cells that had stably integrated and expressed the viral genome. Lysates of two DOL⁺-M-TdT-infected PD31 cell lines showed no significant activity, whereas all three of the DOL⁺-HM-TdT-infected PD31 cell lines tested showed TdT activity levels comparable to that of RL δ 11. Thus, DOL⁺-HM-TdT can direct synthesis of physiological TdT levels stably in PD31 cells.

Introduction of an artificial gene rearrangement substrate into TdT-positive and -negative PD31 cells. An artificial gene rearrangement substrate, DGR, which could be efficiently introduced into lymphocytes was constructed in the DOL retroviral vector (17a) in a manner similar to that previously described (21) (Fig. 1B). On the 3' side of the 5' long terminal repeat was placed the V κ 21-C gene with its flanking heptamer-nonamer sequences followed by the bacterial guanine-xanthine phosphoribosyltransferase gene (*gpt*) gene. Both genes were placed in the opposite transcriptional orientation from the viral long terminal repeat such that the full-length viral transcript could not encode *gpt*. On the 3' side of the *gpt* gene were placed two identical copies of the J κ 1 gene (the duplication of the J κ 1 gene segment relates to other uses of the DGR vector). Immunoglobulin recombinase-mediated inversion of the DNA containing the V κ 21-C and *gpt* genes can join the V κ 21-C gene to either of the two J κ 1s, placing the *gpt* gene in the appropriate orientation and position to allow the proviral transcript to encode *gpt* protein, which can be measured as resistance to mycophenolic acid (Fig. 1C). Selection in mycophenolic acid of a

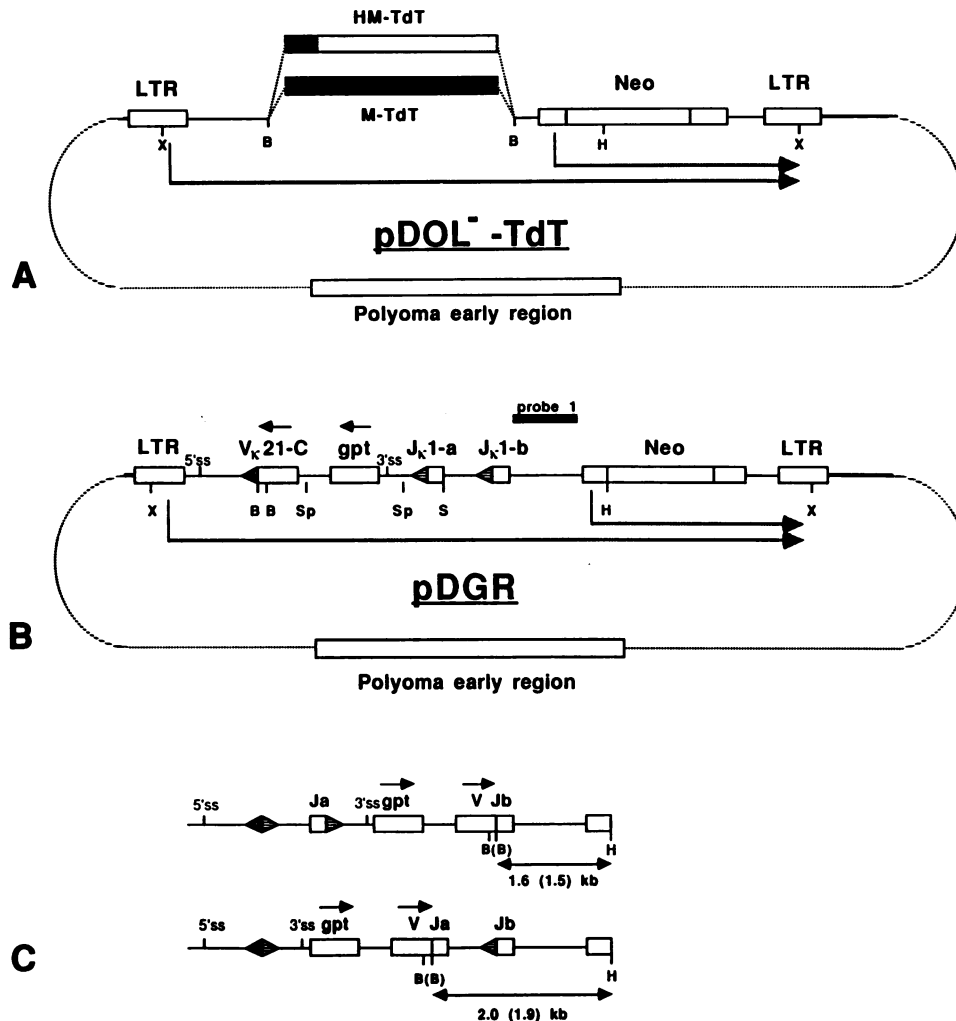


FIG. 1. Structure of the DOL⁻-TdT and DGR retroviral vectors. (A) DOL⁻-TdT retroviral expression vector containing murine leukemia virus long terminal repeats and a packaging sequence to allow for transmission of the vector as a retrovirus. The proviral form of the vector was expected to direct transcription of two RNAs (arrows): one full-length transcript originating at the 5' long terminal repeat promoter and serving as a template for synthesis of TdT and a second transcript originating in the simian virus 40 promoter and encoding the Tn5 neomycin phosphotransferase (*neo*) gene allowing selection of murine cells and bacteria harboring the vector. The pBR322 origin allowed replication of the plasmid form of the vector in bacteria; the polyomavirus early region included in the plasmid form served to increase the titer of retrovirus during transient transfection of the helper virus-free packaging cell line, ψ -2. (B) DGR-directed transcripts (arrows) and conserved heptamer-nonamer sequences (\blacktriangleleft). (C) Structure of the DGR vector immunoglobulin genes after inversional gene rearrangement. The direction of transcription is indicated by arrows. The sizes of *Bam*HI-*Hind*III fragments expected are indicated. Abbreviations: 5'ss, 5' splice site; 3'ss, 3' splice site; LTR, long terminal repeat; B, *Bam*HI; H, *Hind*III; Sp, *Sph*I; X *Xba*I.

population of cells infected with DGR should therefore only permit growth of cells harboring a rearranged DGR provirus while eliminating cells harboring no provirus or an unrearranged provirus. DGR virus was produced by transfecting the ψ -2 packaging cell line with pDGR yielding the producer cell line ψ -iGR (see Materials and Methods).

PD31 cells were infected with the DGR virus by cocultivation with ψ -iGR. The ψ -iGR cell line contained a single copy of an intact DGR genome and did not rearrange the DGR virus with efficiencies detectable in these experiments (data not shown). After 3 days of cocultivation, the PD31 cells were removed from the fibroblast monolayer by gentle aspiration and cultured in the presence of G418 to select for infected cells followed by further selection for *gpt* expression in mycophenolic acid-containing medium. A second group of PD31 cells was infected with DOL⁻-HM-TdT virus and selected in G418; the surviving cells were expanded for

2 weeks, at which time they were infected with DGR virus. Infection of these cells with the DGR virus could not be selected in G418 because the presence of the DOL⁻-HM-TdT provirus rendered both the DGR-infected and uninfected PD31 cells resistant to G418. Instead, the cells were separated into 24 populations to insure independence of the rearrangement events and selected directly in mycophenolic acid-containing medium, according to the strategy of Lewis et al. (21). This treatment selected for cells containing a DGR provirus and which had, in addition, rearranged the DGR provirus allowing for expression of mycophenolic acid resistance. After 2 weeks of selection the vast majority of cells had died and been replaced by resistant cells. Approximately 80% of the wells showed growth.

The configuration of the V_K21-C gene and the J_K genes of DGR virus-infected PD31 cells was assessed by cleavage of DNA with *Bam*HI plus *Hind*III. The cleaved DNA was size

TABLE 1. Expression of TdT activity in fibroblasts and lymphocytes infected with TdT retroviral expression vectors

Vector ^a	Cell line	U ^b
DOL ⁻ -M-TdT	3T3-1a	0.10
DOL ⁻ -M-TdT	3T3-2a	0.12
DOL ⁻ -HM-TdT	3T3-1	0.52
DOL ⁻ -HM-TdT	3T3-2	0.42
DOL ⁻ -HM-TdT	3T3-3	0.26
DOL ⁻ -HM-TdT	3T3-5	0.21
DOL ⁻ -M-TdT	PD31-1a	0.06
DOL ⁻ -M-TdT	PD31-2a	0.11
DOL ⁻ -HM-TdT	PD31-4	1.30
DOL ⁻ -HM-TdT	PD31-6	0.52
DOL ⁻ -HM-TdT	PD31-7	0.52
DOL ⁻ -HM-TdT	RLδ-11	0.45

^a Stably infected cell lines were established by infection with virus produced by ψ -2 producer cell lines, which produced virus at titer of 2×10^4 CFU/ml and contained a single, intact proviral copy.

^b Assays were performed as described in Materials and Methods. Background was about 0.1 U. One unit of enzyme activity is equal to 1 nmol of dGMP incorporated per h.

fractionated by agarose gel electrophoresis, transferred to a nitrocellulose filter, and hybridized to a radioactive probe derived from pBR322 sequences lying 3' to J κ 1-b in DGR (Fig. 1B, probe 1). *Bam*HI cleaved the unrearranged provirus at two closely spaced sites near the 3' end of the V κ 21-C gene, whereas *Hind*III cleaved at a single site in the *neo* gene, releasing a 3.8-kb *Bam*HI-*Hind*III fragment (Fig. 1B). Rearrangement of the DGR V κ 21-C gene to J κ 1-a or to J κ 1-b released 1.9- or 1.5-kb *Bam*HI-*Hind*III fragments, respectively. DGR proviruses of several of the infected cell lines showed the expected 1.9-kb (Fig. 2, lanes 3, 10, and 12) and 1.5-kb (Fig. 2, lanes 4 through 8) fragments. None of the cell lines showed the 3.8-kb fragment indicative of the unrearranged DGR provirus. These results showed that the DGR provirus had rearranged in the PD31 cells such that V κ 21-C had joined to one or the other of the two J κ 1 genes.

Four fragments were observed whose sizes were close to but different from those expected. Two of the fragments could be explained by loss of the *Bam*HI site closest to the V-J junction due to the deletion of nucleotides in this region which accompanies gene rearrangement. In this case, *Bam*HI would cleave only at the *Bam*HI site located 100 base pairs further upstream in the V κ 21-C gene, explaining the observation of fragments of 2.0 kb (data not shown) and 1.6 kb (Fig. 2, lane 12) instead of 1.9 and 1.5 kb. A third fragment of 1.8 kb was seen in several of the cell lines (Fig. 2, lanes 1, 4 through 9, and 11). DNA sequencing showed that this resulted from an irrelevant deletion of approximately 100 base pairs due to a splicing artifact. This splice presumably occurred before packaging of the DGR retroviral DNA in the ψ -iGR producer cell line. A fourth fragment of 1.4 kb probably arose by a splice analogous to that of the 1.8-kb fragment (data not shown).

The presence of more than one of the rearranged fragments in some of the cell lines (Fig. 2, lanes 4 through 8 and 12) was attributed to the presence of two or more cell clones in some of the cell lines. This result was expected, because mycophenolic acid selection of the DGR-infected PD31 cells showed growth in a large fraction of the wells. Each of the cell lines was monoclonal or biclonal, with the exception of cell line 6.2, which was triclonal (data not shown). Each of the DGR proviruses was the result of an independent inte-

gration of the retrovirus as shown by flanking restriction enzyme cleavage sites.

Rescue of the DGR provirus from the PD31 infectants. To determine the sequence of a large number of the V κ 21-C-J κ 1 junctions in the rearranged DGR proviruses, we developed a method for rapidly isolating proviral DNA from the infected PD31 cells lines. Extensive attempts to rescue the proviruses of the DGR-infected PD31 cells by COS cell fusion (7) resulted in an unacceptably low efficiency of rescue (data not shown). Instead, we attempted to rescue the proviral DNA directly from high-molecular-weight DNA isolated from the DGR-infected PD31 cell lines. This procedure took advantage of the ability of the DOL vector to replicate as a plasmid and to express the *neo* gene in *E. coli*, thereby allowing selection of plasmid-containing bacteria with kanamycin. High-molecular-weight DNA from each of the DGR-infected PD31 cell lines was digested to completion with *Xba*I to release the proviral DNA. To enrich for proviral sequences, the DNA was size fractionated by agarose gel electrophoresis, and restriction fragments in the range of 7 to 8 kb were isolated. This size range should include proviruses with each of the observed rearrangements. The isolated DNA was circularized and used to transform *E. coli*.

Transformation of *E. coli* without DNA or with 3T3 DNA processed as described above resulted in no colonies. Transformation with DNA from the DGR-infected PD31 cell lines resulted in an average of 20 bacterial colonies per cell line. Plasmid DNA was isolated from six colonies of each transformation. The DNA was cleaved with *Bam*HI and *Hind*III and size fractionated by agarose gel electrophoresis. The sizes of the *Bam*HI-*Hind*III fragments containing the V κ 21-C-J κ 1 junction were compared with the sizes of the fragments observed with the same restriction digest in the genomic DNA of the DGR-infected PD31 cell lines. In each case the *Bam*HI-*Hind*III fragment observed in the genomic DNA was of the same size as that of the rescued plasmid (data not shown). In cell lines that showed two or three different-sized *Bam*HI-*Hind*III fragments, the rescued plasmid DNAs showed plasmids with fragments of the corre-

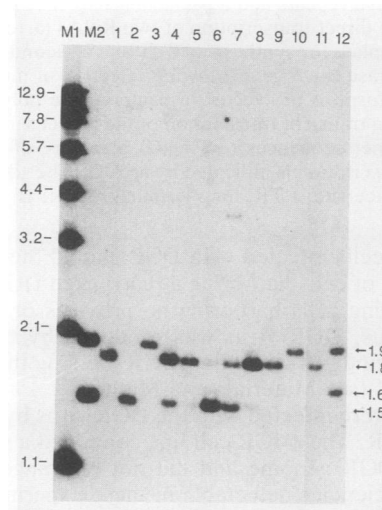


FIG. 2. Rearrangements of DGR provirus in infected PD31 cells. High-molecular-weight DNA (10 μ g) was digested with *Bam*HI-*Hind*III, electrophoresed through 1.2% agarose, transferred to nitrocellulose, and hybridized to radioactive probe 1 DNA. Lanes: M1, pBR322 markers; M2, markers derived from rescued plasmids; 1 through 12, cell lines 6.16, 6.1, 6.5, 6.9, 6.10, 6.11, 6.14, 6.15, 6.16, 6.18, 6.12, and 6.4, respectively.

TABLE 2. Sequences around V-J junction of rearranged DGR rescued plasmids

GGATCCTCC ^a	GTGGACGTTTCGGTGGAGGCACCAAG ^b	Cell line ^c	Fragment size ^d	TdT virus	TdT level ^e
----- AC -----	-----	10.1	1.5		ND
-----	-----	10.2	1.5		NT
-----	-----	10.3	1.5		ND
-----	-----	10.4	1.9		ND
-----	-----	10.5	1.5		NT
-----	-----	10.6	1.9		NT
-----	-----	10.7	1.9		NT
-----	-----	10.8	2.0		NT
-----	-----	6.2	1.6	DOL ⁻ -HM-TdT	NT
-----	-----	6.2	2.0	DOL ⁻ -HM-TdT	NT
----- GA -----	-----	6.2	1.9	DOL ⁻ -HM-TdT	NT
-----	-----	6.4	1.4	DOL ⁻ -HM-TdT	21
-----	-----	6.5	1.8	DOL ⁻ -HM-TdT	ND
-----	-----	6.5	1.9	DOL ⁻ -HM-TdT	ND
----- G -----	-----	6.6	1.8	DOL ⁻ -HM-TdT	91
-----	-----	6.7	1.5	DOL ⁻ -HM-TdT	ND
-----	-----	6.8	1.5	DOL ⁻ -HM-TdT	17
-----	-----	6.11	2.0	DOL ⁻ -HM-TdT	NT
-----	-----	6.9	1.5	DOL ⁻ -HM-TdT	30
----- TCCCA -----	-----	6.9	1.9	DOL ⁻ -HM-TdT	30
-----	-----	6.10	1.5	DOL ⁻ -HM-TdT	16
----- CT -----	-----	6.10	1.8	DOL ⁻ -HM-TdT	16
----- GA -----	-----	6.15	1.5	DOL ⁻ -HM-TdT	14
----- CT -----	-----	6.15	1.8	DOL ⁻ -HM-TdT	14
-----	-----	6.16	1.5	DOL ⁻ -HM-TdT	10
----- A -----	-----	6.18	1.5	DOL ⁻ -HM-TdT	100

^{a,b} Sequences of the 3' end of V_K21-C (a) and the 5' end of J_K1 (b) up to the conserved heptamer-nonamer sequences. Dashes indicate germ line sequence. Non-germ line nucleotides are shown in boldface type.

^c PD31 cell lines containing independent insertions of the DGR provirus.

^d HindIII-BamHI fragment size in kilobases of rescued plasmid.

^e TdT was assayed in control cell lines by RNA blotting with a radioactive M-TdT probe. DOL⁻-HM-TdT virus-infected cell lines were assayed for TdT activity as described in Materials and Methods. Units are arbitrary and set to 100 for the most active cell line. This cell line had activity equivalent to that of RL δ 11. ND, Not detectable; NT, not tested.

sponding sizes. Thus, it was possible by this method to rapidly and efficiently rescue single-copy or low-copy-number plasmids directly from the genomic DNA of cells. The rescued plasmids accurately reflected the proviruses contained in the DGR virus-infected PD31 cell lines.

Sequences of the V_K21-C-J_K1 junctions of the rescued plasmids. The nucleotide sequence of V_K21-C-J_K1 junctions of the rescued plasmids was determined by labeling the BamHI site of the plasmid DNA and subjecting the DNA to the sequencing reactions of Maxam and Gilbert (26). Approximately 100 base pairs of sequence was determined for each fragment. Of the sequences derived from the control cell lines, all except for one of the eight V-J junctions contained direct joins of V_K21-C to J_K1-a or to J_K1-b with, as expected, loss of a few nucleotides (22) (Table 2). Of the plasmids derived from DOL⁻-HM-TdT virus-infected PD31 cells, 39% (7 of 18) showed non-germ line nucleotides at the V_K21-C-J_K1 junction (Table 2). All of the N regions consisted of 1 or 2 base pairs, except for one 5-base-pair N region. Of the nucleotides further 3' of the V-J junction, all sequenced nucleotides corresponded to the sequence of J_K1, indicating that the non-germ line-encoded nucleotides were not simply the result of a high mutation rate of nucleotides in DGR.

Because each DGR virus-infected PD31 cell clone contained an independent integration of the DOL⁻-HM-TdT virus, it was possible that the cell lines expressed different levels of TdT. In addition it was desirable to test the TdT level in the control cell line that showed an N region to test for the presence of low-level TdT expression. The results of TdT assays of lysates prepared from several of the DOL⁻-HM-TdT-infected cell lines showed that three of the cell

lines had TdT levels that were not detectable over background (Table 2). These cell lines showed germ line V-J junctions. One of the cell lines contained two different DGR rearrangements. Of the rearrangements which showed N-region addition, all (five of five) of the corresponding cell lines showed significant TdT levels. The TdT in the cell lines varied considerably from physiological to undetectable levels, probably due to differences in the integration site of the DOL-HM-TdT provirus in each cell line. To determine as sensitively as possible the TdT levels in the control cell line, 10.1, whose rescued plasmid showed an N region, RNA hybridization analysis was performed. Hybridization of a radioactive TdT probe to size-fractionated RNA from this cell line showed no detectable TdT RNA, whereas RL δ 11 showed a strong band (data not shown). The 10.1 cell line therefore contained at most a very low level of TdT.

DISCUSSION

We have investigated the role of TdT in the insertion of N regions into the junctional regions of immunoglobulin genes. Previous studies on the role of TdT in the insertion of N regions relied on cell lines which differed in TdT level but which also could have differed in the expression of gene products which might be coordinately regulated with the TdT gene (11, 39). In the present study we sought to overcome this difficulty by establishing cell lines which differed from one another only by the expression of TdT. This was accomplished by the introduction of a TdT retroviral expression vector into a TdT⁻ cell line. The resulting fibroblast and pre-B-cell lines showed physiological TdT

activity levels. TdT expression in the infected cell lines was encoded by the introduced TdT gene and not the endogenous TdT gene, because introduction of a control retroviral expression vector containing a mutant TdT cDNA did not induce TdT expression. It was of interest that the presence of TdT had no obvious toxic effect on 3T3 cells, although it could lead to subtle effects such as an increased mutation rate.

Comparison of immunoglobulin V and J sequences rearranged in the presence or absence of TdT showed a striking difference. About 10% of the control segments contained N regions, whereas 50% of the segments rearranged in the presence of detectable TdT contained N regions. The control value (one of eight) is consistent with a previous observation of one of nine rearrangements containing N regions in the absence of detectable TdT with a gene rearrangement substrate similar to that used here (22). The cause of these apparently TdT-independent N regions is not clear. They could be caused by a very low level of TdT, undetectable even by mRNA analysis; by transient expression of TdT in a rare cell, which would not later be detected; or by a truly TdT-independent process. Short runs of sequence have been found in nonimmunoglobulin DNA joints (33), and the mechanism behind our observation could be similar.

Although 7 of 18 genes rearranged in HM-TdT virus-infected cells contained N regions, at least four of those lacking N regions were found to have occurred in cells that lacked TdT expression. These cells contained the integrated HM-TdT provirus but, presumably due to position effects or promoter competition, did not express a detectable amount of enzyme. Thus, at least 50% of the rearrangements made in the presence of TdT had N regions.

Our determinations of N-region lengths are minimum estimates because we are assuming that every germ line-encoded nucleotide found represents a true germ line nucleotide. Some fraction certainly represents N-region nucleotides that fortuitously agreed with the germ line sequence. This problem should affect the control and experimental data similarly.

We conclude that TdT can increase the frequency of N region insertion by at least fivefold. Actually, the control level may represent a process quite different from that catalyzed by TdT, but we presently have no way to distinguish the two possibilities. Furthermore, the frequency of N-region insertion appears low compared with that in authentic heavy-chain joints. This may reflect the use of an artificial substrate or the use of κ -gene segments, which ordinarily probably do not rearrange in the presence of TdT. We chose to investigate κ -gene rearrangements purposely to illustrate the efficiency of TdT, but we may have given the enzyme a particularly difficult task. This may also explain the relatively short length of most of the N regions we observed.

The results presented here are likely to apply equally well to T-cell receptor N regions; both gene families use the same recombinase (39) and the same conserved heptamer-nonamer sequence (9) and rearrange at a stage of lymphocyte maturation during which TdT is present (18). N-region diversity may take on added importance in the generation of the T-cell receptor repertoire in light of the apparent lack of somatic hypermutation of T-cell receptor genes (9), in contrast to immunoglobulin genes where such mutations play a large role in diversification (16). The lack of somatic hypermutation may reflect the need to avoid modification in the periphery of the precise joint recognition of major histocompatibility product plus antigen that is established

during thymic selection. The joint recognition process may even require that germ line sequence be maintained in T-cell receptor genes outside of the V-J junction region.

ACKNOWLEDGMENTS

This work was supported by a grant from the American Cancer Society. David G. Schatz is a recipient of a fellowship from the Life and Health Insurance Medical Research Fund.

M.R. was on leave from Biogen, Cambridge, Mass.

LITERATURE CITED

- Alt, F., and D. Baltimore. 1982. Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-J_H fusions. *Proc. Natl. Acad. Sci. USA* **79**:4118-4122.
- Baltimore, D. 1974. Is terminal deoxynucleotidyl transferase a somatic mutagen in lymphocytes? *Nature (London)* **248**:409-410.
- Barth, R. K., B. S. Kim, N. C. Lan, T. Junkapiller, N. Sobiek, A. Winoto, H. Gershenfeld, C. Okada, D. Hansburg, I. L. Weissman, and L. Hood. 1985. The murine T-cell receptor uses a limited repertoire of expressed V β gene segments. *Nature (London)* **316**:517-523.
- Blackwell, T., and F. Alt. 1984. Site-specific recombination between immunoglobulin D and J_H segments that were introduced into the genome of murine pre-B cell line. *Cell* **37**:105-112.
- Bollum, F. J. 1974. Terminal deoxynucleotidyl transferase, p. 145. *In* P. D. Boyer (ed.), *The enzymes*, vol. 10. Academic Press, Inc., New York.
- Bollum, F. J., and M. Brown. 1979. A high molecular weight form of terminal deoxynucleotidyl transferase. *Nature (London)* **278**:191-192.
- Cepko, C. L., B. E. Roberts, and R. C. Mulligan. 1984. Construction and application of a highly transmissible murine retrovirus shuttle vector. *Cell* **37**:1053-1062.
- Chang, L. M. S. 1971. Development of terminal deoxynucleotidyl transferase activity in embryonic calf thymus gland. *Biochem. Biophys. Res. Commun.* **44**:124-128.
- Chien, Y. H., N. R. J. Gascoigne, J. Kavalier, N. Lee, and M. M. Davis. 1984. Somatic recombination in a murine T-cell receptor gene. *Nature (London)* **309**:322-326.
- Coleman, M. S., J. J. Hutton, P. de Simone, and F. J. Bollum. 1974. Terminal deoxyribonucleotidyl transferase in human leukemia. *Proc. Natl. Acad. Sci. USA* **71**:4404-4408.
- Desiderio, S., G. Yancopoulos, M. Paskind, E. Thomas, M. Boss, N. Landau, F. Alt, and D. Baltimore. 1984. Insertion of N-regions into heavy-chain genes is correlated with expression of terminal deoxynucleotidyl transferase in B-cells. *Nature (London)* **311**:752-755.
- Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction fragments to high specific activity. *Anal. Biochem.* **132**:6-13.
- Gearhart, P. J., N. D. Johnson, R. Douglas, and L. Hood. 1981. IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. *Nature (London)* **291**:29-34.
- Graham, F. L., and A. J. van der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**:456-467.
- Kato, K., H. M. Goncalves, G. E. Gouts, and F. J. Bollum. 1967. Deoxynucleotide-polymerizing enzymes of calf thymus gland. *J. Biol. Chem.* **242**:2780-2789.
- Kim, S., M. Davis, E. Sinn, P. Patten, and L. Hood. 1981. Antibody diversity: somatic hypermutation of rearranged V_H genes. *Cell* **27**:573-581.
- Koiwai, O., T. Yokota, T. Kageyama, T. Hirose, S. Yoshida, and K. Arai. 1986. Isolation and characterization of bovine and mouse terminal deoxynucleotidyltransferase cDNA's expressible in mammalian cells. *Nucleic Acid Res.* **14**:5777-5781.
- Korman, A. J., J. D. Frantz, J. L. Strominger, and R. C. Mulligan. 1987. Expression of human class II major histocompatibility complex antigens using retrovirus vectors. *Proc. Natl.*

- Acad. Sci. USA **84**:2150–2154.
18. Kung, P. C., A. E. Silverstone, R. P. McCaffrey, and D. Baltimore. 1976. Murine terminal deoxynucleotidyl transferase: cellular distribution and response to cortisone. *J. Exp. Med.* **141**:855–865.
 19. Kurosawa, Y., and S. Tonegawa. 1982. Organization, structure, and assembly of immunoglobulin heavy chain diversity DNA segments. *J. Exp. Med.* **155**:201–218.
 20. Landau, N. R., T. P. St. John, I. L. Weissman, S. C. Wolf, A. E. Silverstone, and D. Baltimore. 1984. Cloning of terminal transferase cDNA by antibody screening. *Proc. Natl. Acad. Sci. USA* **81**:5836–5840.
 21. Lewis, S., A. Gifford, and D. Baltimore. 1984. Joining of V_κ to J_κ gene segments in a retroviral vector introduced in lymphoid cells. *Nature (London)* **308**:425–428.
 22. Lewis, S., A. Gifford, and D. Baltimore. 1985. DNA Elements are asymmetrically joined during the site-specific recombination of κ immunoglobulin genes. *Science* **228**:677–685.
 23. Lewis, S., N. Rosenberg, F. Alt, and D. Baltimore. 1982. Continuing κ gene rearrangement in a cell line transformed by Abelson murine leukemia virus. *Cell* **30**:807–816.
 24. Mann, R., R. Mulligan, and D. Baltimore. 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* **33**:153–159.
 25. Max, E. E., J. V. Maizel, and P. Leder. 1981. The Nucleotide sequence of a 5.5-kilobase DNA segment containing the mouse κ immunoglobulin J and C region genes. *J. Biol. Chem.* **256**:5116–5120.
 26. Maxam, A. M., and W. Gilbert. 1980. Sequencing end labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**:499–560.
 27. McCaffrey, R., D. F. Smoler, and D. Baltimore. 1973. Terminal deoxynucleotidyl transferase in a case of childhood acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. USA* **70**:36–40.
 28. McCaffrey, R. P., T. A. Harrison, R. Parkman, and D. Baltimore. 1975. Terminal deoxynucleotidyl transferase activity in human leukemia cells and in normal human thymocytes. *N. Engl. J. Med.* **292**:775–780.
 29. Nakamura, H., K. Tanabe, S. Toshida, and T. Morita. 1981. Terminal deoxynucleotidyltransferase of 60,000 daltons from mouse, rat, and calf thymus. *J. Biol. Chem.* **256**:8745–8751.
 30. Parker, B. A., and G. R. Stark. 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. *J. Virol.* **31**:360–369.
 31. Peterson, R. C., L. C. Cheung, R. J. Mattaliano, L. M. S. Chang, and F. J. Bollum. 1984. Molecular cloning of human terminal deoxynucleotidyl transferase. *Proc. Natl. Acad. Sci. USA* **81**:4363–4367.
 32. Peterson, R. C., L. C. Cheung, R. J. Mattaliano, S. T. White, L. M. S. Chang, and F. J. Bollum. 1985. Expression of human terminal deoxynucleotidyltransferase in *Escherichia coli*. *J. Biol. Chem.* **260**:10495–10502.
 33. Roth, D. B., T. N. Porter, and J. H. Wilson. 1985. Mechanism of nonhomologous recombination in mammalian cells. *Mol. Cell. Biol.* **5**:2599–2607.
 34. Sakano, H., K. Huppi, G. Heinrich, and S. Tonegawa. 1979. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature (London)* **280**:288–294.
 35. Silverstone, A. E., H. Cantor, G. Goldstein, and D. Baltimore. 1976. Terminal deoxynucleotidyl transferase is found in prothymocytes. *J. Exp. Med.* **144**:544–550.
 36. Silverstone, A., L. Sun, O. N. Witte, and D. Baltimore. 1980. Biosynthesis of murine terminal deoxynucleotidyl transferase. *J. Biol. Chem.* **255**:791–795.
 37. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503–517.
 38. Vogelstein, B., and D. Gillespie. 1979. Preparation and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci. USA* **76**:615–619.
 39. Yancopoulos, G. D., K. T. Blackwell, H. Suh, L. Hood, and F. W. Alt. 1986. Introduced T cell receptor variable region gene segments recombine in pre-B cells: evidence that B and T cells use a common recombinase. *Cell* **44**:251–259.